Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 10 (2012) 952 – 959

Procedia Computer Science

The 2nd International Symposium on Frontiers in Ambient and Mobile Systems (FAMS)

Emerging URL Patterns in Mobile Websites: A Preliminary Results

Atheer S. Al-Khalifa[a] and Hend S. Al-Khalifa[b]*

[a]*Electronic and Computer Institute, King Abdulaziz City of Science and Technology, Riyadh, Saudi Arabia*
[b]*Information Technology Department, King Saud University, Riyadh, Saudi Arabia*

Abstract

In the last couple of years, Mobile Web has emerged as a concept that presents the "web on the go" with dramatic development in its size and technologies. The first entry point for any mobile website is through its Uniform Resource Locator (URL). A URL plays a major part of any website interface. With appropriate URL addresses, websites can be located and remembered easily.
In this paper we first discuss the different approaches used by desktop websites regarding the construction of their mobile websites' URLs. Then we present the results of manually analyzing a set of 100 mobile websites to identify and distill the most commonly used URL patterns in these websites.

Keywords: Mobile web, URL Patterns, Mobile browsing, domain schema, usability

1. Introduction

In the past few years, Smart phones industry showed an incredible growth in terms of owners' number, accompanied by the increasing availability of wireless networks and large bandwidth of data transmission speeds e.g. 3G/4G networks. As a consequence of this rapid development, browsing mobile websites had gained an immerse importance to the point where predictions of its internet usage will exceed the desktop counterpart by 2015 [1].

Mobile web is not limited to Smart phones only but exceeds to target tablets and portable mp3 players that support internet access, which results in an increased number of users accessing the internet on the go.

The size of mobile web has not been statistically determined yet, however resorting to Google search engine we can get a rough estimate of the number of mobile websites. By searching for the pattern

* Corresponding author. Tel.: +966-1-4743090
E-mail address: hendk@ksu.edu.sa.

1877-0509
doi:10.1016/j.procs.2012.06.130

"allinurl: *.mobi", Google shows that the number of mobile websites with the domain ".mobi" is approximately 1650 Million websites , which is considered a huge number for a recent trend.

Users differ in their methods of reaching mobile websites. Some prefer relying on hyperlinks provided by different referral resources, e.g. messaging applications. Others use search engines to locate a website by entering its name as a search term. However, some just type the desktop website's URL and leave it to the website to redirect the mobile browser to the appropriate version.

Taking care of URL schema, i.e. the way URLs are named and constructed, will help users locate websites easily and intuitively [2]. This recommendation also applies to the URLs of mobile websites. In order to understand the way mobile websites' URLs are constructed, we posed the following research questions:

1. What are the most recurring URL patterns of mobile websites?
2. Do mobile websites' URL patterns differ based on their genre?

In this paper we present the results of manually analyzing the URLs of 100 mobile websites based on their genre, and observing the different methods used in designing URLs in order to distill the most commonly used URL patterns in mobile websites.

2. Background

A URL is a sequence of characters that identifies a resource on the internet. The Internet Official Protocol Standards (STD 1) presented the URL as one of the methods for accessing resources by providing "abstract identification of its location" [3]. Once a resource is located, a system can perform different operations on it [3].

Users often reach desktop websites through mechanisms that involve exposure to URLs via direct access or referrals. This is also true in case of mobile devices, where users interact with URLs using two approaches. The first is by referral, where a user follows a hyperlink shared by social network applications, e-mail client, SMS & instant messaging applications or other web pages. It also can be obtained by scanning a 2D barcode (e.g. Quick Response or high capacity color barcodes)[a] or real world elements for an augmented reality application[b]. Moreover, users use search engines to find a specific website either by entering a website's name or typing related keywords.

The second approach is by typing the URL directly, which is usually for an instant necessity of information in a website. Users prefer the former way because typing on screen-limited devices is difficult and chances of entering the wrong URL are high. Thus, the W3C mobile web document mentioned that lengthy URLs which contain much punctuation are difficult for users to enter correctly [4].

In mobile web, the relation between the URL of a desktop website and its mobile equivalent is dynamic. Many desktop websites have dealt differently with assigning URLs to their mobile compatible versions. Figure 1 summarizes the four approaches for constructing mobile URLs.

---

[a] What are 2D barcodes: http://tag.microsoft.com/what-is-tag/2d-barcodes.aspx
[b] Layar Browser – Layar Vision: http://www.layar.com/browser/layar-vision/

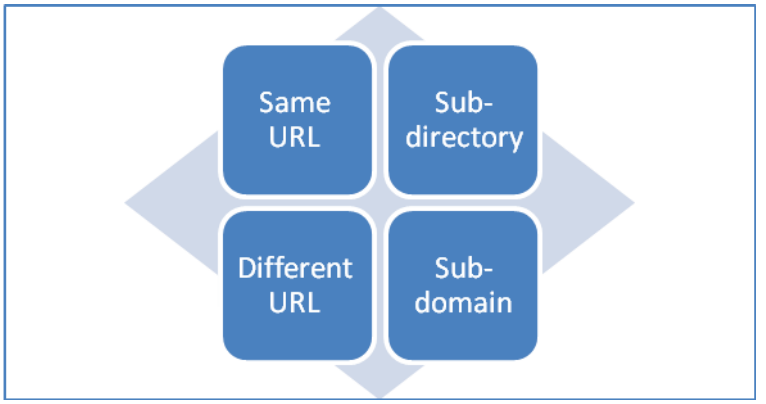

Fig. 1. Four approaches for constructing mobile URLs

Some prefer keeping the same address as the desktop website and having the advantage of saving users from remembering a new address. However, in this case the mobile browser will modify and adapt the content based on its capabilities. One method is to create mobile-optimized style sheet that is applied to the entire website [10].This approach, in general, can be restricting if the website's content is exceedingly complex and dynamic [5].

The second approach is to have a different registered domain. The W3C Mobile Web Initiative (MWI) provided a top level domain specifically to mobile websites (.mobi). Having such domain can help facilitate search engines' recognition of a website as being mobile compatible. But on the other hand, it may be costly to obtain such domain for the same name and in some cases it might be already registered for another owner. Similarly, some websites would have a total different URL address, which is not advisable by W3C because it will involve some overhead in terms of expenses and maintenance. In addition, users will have to remember another address for the website.

The Third approach is to redirect the URL to a subdirectory with mobile compatible layout. Thus, the URL is still pointing to the main website and the redirection is done by the server, e.g. http://example.com/mobile. A main disadvantage of this approach is the possibility of crowding search engines by indexing the mobile website as a redundant content or what is known as "Duplicate URL, Same Text (DUST)" [6].

Similar to the third approach, is to create a specific sub-domain to a website, which is also a common method. This approach will not require any additional costs for buying another domain as well as it helps manage the website easier by having a single domain. As an example: http://mobile.example.com . Both the sub-domain and subdirectories approaches share an advantage, which is benefiting the mobile website from its desktop equivalent. The benefits are evident in terms of: traffic, links and keyword ranking acquisition in search engines' listings, which dismisses the need for the mobile website to start attracting traffic again [6].

In all cases, once a user arrives to a webpage using a mobile device, it will be redirected to the mobile website. The redirection can be done using client side scripting that detects the screen's size and redirects to the appropriate version accordingly [7], or server side scripting by reading the HTTP Header coming from the mobile browser which contains the user agent's strings and looking up this string in pre-compiled database such as WURFL[c]. It then determines whether the received user agent's string is from a

c WURFL is "a Device Description Repository which contains the descriptions of thousands of mobile devices": http://wurfl.sourceforge.net/

mobile device or not [8]. Either way, since redirection requires extra data transfer, it can be costly for users accessing the Internet via 3G/4G networks or prepaid wireless subscription.

3. Method

To answer our previous research questions, a random sample of 100 unique mobile websites' URLs were gathered and classified into categories based on their genre. This small set does not represent the whole population that exceeds 1650 million mobile websites but gives a primitive look into what if a significant portion of the mobile web is analyzed. The samples were then compared to their desktop URL equivalent, indicating whether they were identical or not. From each category we distilled the most used patterns, and computed the frequency of appearances for each pattern in the whole data set.

The mobile websites were classified into nine genres: news sites, social networks, information providers, blogs, gaming websites, corporate websites, web mail, e-commerce and media sharing sites. This classification is based on Wikipedia's article "Website", under "Types of Websites" section [9].

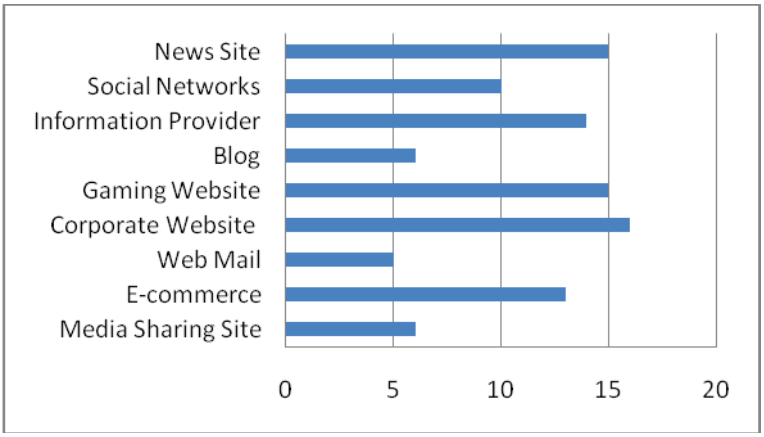

Fig. 2. Breakdown of mobile websites by genre

Figure 2 shows the number of collected mobile websites based on their genres. The presence of quantity distinction between genres can be clearly seen in the figure, wherefore blogs, web mail and media sharing sites have the lowest number of samples compared to other mobile websites due to the following factors: (1) The wide usage of Content Management Systems (CMS) in blogging that adapts the same mechanism for their mobile websites, for instance, WordPress provides a plug-in called Mobile Edition plug-in[d] that helps in mobilizing blogs. (2) As for web mail and media sharing sites, it can be attributed to the magnitude of those categories' websites compared to other types on the Internet.

4. Results

Table 1 provides an overview of websites categories for our data set and their most occurred patterns in terms of appearances per category. It can be seen from the table that "m." sub-domain is the most common pattern used in seven categories out of nine, which shows a general consensus for this approach even with distinct different genres.

---

d http://wordpress.org/extend/plugins/wordpress-mobile-edition/

Table 1. Breakdown of most common URL patterns for each genre with the percentage of occurrence

Category	Common pattern	Percentage of URLs in the genre
News site	m. sub-domain	40%
Social networks	m. sub-domain	30%
Information provider	m. sub-domain	36%
Blog	No difference (same URL)	67%
Gaming website	m. sub-domain	47%
Corporate website	m. sub-domain	44%
Web mail	/(different directory)	60%
E-commerce	m. sub-domain	54%
Media sharing site	m. sub-domain	50%

On the other hand, blog websites prefer keeping the same URL address as their desktop version. This behaviour can be attributed, as mentioned previously, to the common use of CMS's plug-ins in blogging systems. Whereas web mail websites showed their preference toward sub-directories approach, which can be attributed to security reasons. To prevent violation of users' mail boxes, some mail service providers allocate a unique sub-directory for each user [11].

Table 2. Classification of Mobile websites' URL types based on their pattern proportion in the whole data set

Type	Pattern	% of URLs in the pattern	
No difference (same URL)	------	14%	
Prefix (sub-domain)	m.	43%	60%
	mobile.	10%	
	i.	1%	
	touch.	1%	
	iphone.	3%	
	mw.	1%	
	www1.	1%	
Suffix (all domains)	.mobi	1%	
Directory	/mobile	6%	16%
	/m	3%	
	/(different directory)	4%	
	/mob	1%	
	/iphone	1%	
	/pda	1%	
Entirely different domain		4%	

Table 2 shows the types of mobile websites' URLs and their percentages in the whole data set. As shown in the table, the sub-domain prefix "m." is the most used pattern representing 43% of the entire data set. This emphasizes that "m." sub-domain is the easiest to type and remember for mobile users. Secondly comes 14% of websites which prefer keeping their URLs intact while providing a mobile compatible layout. Whereas the "mobile." prefix sub-domain comes third in place representing 10% of occurred patterns in our data set.

Having a different domain or using the suffix ".mobi" are the least desirable approaches compared to other options. This can be attributed to the extra costs required when reserving a new domain.

5. Discussion

Regardless of specific naming, choosing a sub-domain is the major approach used in mobile websites URLs, followed by the usage of directories, which shows the orientation of relating the mobile site to its desktop equivalent. It was also noticed that some mobile websites customize their domains based on the type of the mobile phone, for instance, some provided iphone or touch prefix to indicate specific type of mobile devices. We also presume that these two methods are more preferable due to the necessity of keeping the main domain intact, specifically for Sponsored Top Level Domains (STLD) such as edu, gov, etc.

Even with the diversity of genres selected in this experiment, it can be clearly seen that it did not make a difference in identifying the most common pattern, where seven different genres had selected the same type of pattern. This finding answers our second research question.

To further justify our findings, we searched for the two most widely occurred patterns in our data set (i.e. m sub-domain and mobile sub-domain) using the Google search engine operator "site:". The results are shown in table 3.

Table 3. Google search results for m and mobile sub-domains

Pattern	number of results
m.*	28,700,000
mobile.*	7,280,000

These results are rough, which might contain some anomalies in the returned results, i.e. sites that do not have an m sub-domain. However, the returned results show clearly the huge difference between the two patterns, and surprisingly the ratio for the number of mobile sub-domain to the m sub-domain returned by Google results is nearly equivalent to the ratio of our experiment (1:4). In other words, for each mobile sub-domain there are four times as many as m sub-domain. This marks the m sub-domain as a recurring best practice when deciding the schema of a mobile website URL.

6. Conclusion

In this paper we presented the results of a pilot study that investigated current patterns used in mobile websites URLs. Our results found that the usage of sub-domains as well as directories were among the approaches used in mobile websites URLs. The results also show diversity in URL patterns and lack of standardization. This diversity might be attributed to the fragmentation of mobile device market with different capabilities.

One weakness of this study is in its small data set; however, in our future work we intend to justify our findings by employing a large data set of mobile websites and automating the process of URL pattern identification.

References

[1] The Growth of Mobile Marketing and Tagging, http://tag.microsoft.com/community/blog/t/the_growth_of_mobile_marketing_and_tagging.aspx

[2] URL as UI, http://www.useit.com/alertbox/990321.html

[3] Resource Locators (URL), http://www.hjp.at/doc/rfc/rfc1738.html

[4] Mobile Web Best Practices 1.0, http://www.w3.org/TR/mobile-bp/

[5] Optimizing a website for mobile devices, http://modernnomads.info/wiki/index.php?page=Optimizing+a+website+for+mobile+devices

[6] Krum, C.: Mobile Marketing. Pearson Education, Indianapolis (2010)

[7] Find mobile visitors to your website and redirect them to your mobile site, http://justtechnika.com/articles/find-mobile-visitors-to-your-website-and-redirect-them-to-your-mobile-site/

[8] How to build your mobile site. Incentivated, London (2011), http://www.incentivated.com/whitepaper/mobileinternet/

[9] Website. Wikipedia, http://en.wikipedia.org/wiki/Website#Types_of_websites

[10] Iglesias, E. and Meesangnil, W. (2011), Mobile website development: From site to app. Bulletin of the American Society for Information Science and Technology, 38: 18–23. doi: 10.1002/bult.2011.1720380108

[11] A. Singh, B. Singh, and H. Joseph, "Vulnerability Analysis for Mail Protocols," in Vulnerability Analysis and Defense for the Internet, vol. 37, A. Singh, B. Singh, and H. Joseph, Eds. Boston, MA: Springer US, pp. 47–70.